

# Spectral density of products of Wishart dilute random matrices. Part I: the dense case

**Thomas Duplic**

Department of Mathematics, King's College London, Strand, London WC2R 2LS,  
United Kingdom

**Isaac Pérez Castillo**

Department of Mathematics, King's College London, Strand, London WC2R 2LS,  
United Kingdom

Departamento de Sistemas Complejos, Instituto de Física, UNAM, P.O. Box 20-364,  
01000 México D.F., México

**Abstract.** In this work we study the spectral density of products of Wishart diluted random matrices of the form  $X(1) \cdots X(M)(X(1) \cdots X(M))^T$  using the Edwards-Jones trick to map this problem into a system of interacting particles with random couplings on a multipartite graph. We apply the cavity method to obtain recursive relations in typical instances from which to obtain the spectral density. As this problem is fairly rich, we start by reporting in part I a lengthy analysis for the case of dense matrices. Here we derive that the spectral density is a solution of a polynomial equation of degree  $M + 1$  and obtain exact expressions of it for  $M = 1, 2$  and  $3$ . For general  $M$ , we are able to find the exact expression of the spectral density only when all the matrices  $X(t)$  for  $t = 1, \dots, M$  are square. We also make some observations for general  $M$ , based admittedly on some weak numerical evidence, which we expect to be correct.

## 1. Introduction

We can safely say that research in random matrix theory is alive and well. This is not unsurprising considering the ever expanding range of applications from Mathematics, to Physics, to Economics, to Biology, to Engineering, and so on. The field is expanding so fast that it is becoming more and more difficult to be up to date or even trying to master all mathematical approaches available (SuSy, determinantal, spin glass techniques, free probability theory, etc).

From the various active research topics, there is a visibly recent and growing interest in understanding the statistical properties of products of random matrices [1, 2, 3, 4, 5, 6, 7] even though this topic had already been moderately active in the past [8]. While a sizeable part of these studies deal with complex matrices, the purpose of this series of works (parts I and II) is to complement the already existing knowledge by extending the current studies to the realm of real and diluted random matrices ‡. As it has been noticed in several occasions, such a simple imposition in the matrix entries (i.e. many zero entries proportional to the matrix size and few non-zero entries independent of the matrix size) complicates the mathematical analysis. In fact, the mathematical derivations have become so voluminous in the preparation of this work that we have preferred to split it into two parts: The first part corresponds to the analysis of the ensemble of matrices in the dense limit, while the second part corresponds to the analysis of general diluted matrices, while paying particular attention on generalising the results of part I to the case of random regular graphs. We have decided to start with dense matrices for two reasons: firstly, it seems relatively “the easiest case”; secondly, there are already quite some known results in this case -derived by different methods-. We expect that part I will be a nice introduction to the spin-glass methods we use and, in turn, will complement the existing alternative derivations (see for instance [3, 1, 2] and references therein), hopefully bringing new results as well.

This paper is organized as follows: in section 2 we introduce the problem mathematically and see how the spectral density can be derived from the statistical properties encoded in a partition function. In section 3 we tackle the effective Hamiltonian by using the cavity method and perform the dense limit. We also show that the spectral density is given by one of the solutions of a polynomial equation of degree  $M + 1$ . In section 4 we consider the cases for  $M = 1, 2$ , and  $3$ , which corresponds to those instances in which the polynomial equation can be solved by radicals, and compare our exact results with numerical diagonalisation. Section 5 is a mixed bag of exact results and numerical observations that lead to, what we believe are, exact results. The last section 6 is for conclusions and advancing what is to be expected in part II of this technical report.

‡ A good summary of these results was already presented by T. Dupic in a poster at the *Mini-Conference on Statistical Mechanics of Glassy and Disordered Systems*, King’s College London, 20-21 May 2013.

## 2. Spectral density as a spin glass problem

While there exist various mathematical approaches to study statistical properties of matrix ensembles, as soon as these matrices are diluted (i.e. many zero entries and a few non-zero entries) our choice of available techniques rapidly narrows. Unfortunately, this is the situation we encounter here, so we take the safe route of understanding the spectral density as a spin glass problem and applying the well-understood techniques of the latter research area to say something about the original problem.

Let us therefore consider the following ensemble of matrices: given  $M$  matrices  $X(t)$  for  $t = 1, \dots, M$ , we define the matrix product:

$$Y = X(1) \cdots X(M)(X(1) \cdots X(M))^T. \quad (1)$$

Here each matrix  $X(t)$  is a  $N(t) \times N(t+1)$  rectangular matrix. This implies that  $Y$  is a positive-definite symmetric matrix of order  $N(1) \times N(1)$ . Denoting as  $\{\lambda_i^Y\}_{i=1, \dots, N(1)}$  its collection of real and positive eigenvalues, its spectral density is defined by

$$\rho_Y(\lambda) = \frac{1}{N(1)} \sum_{i=1}^{N(1)} \delta(\lambda - \lambda_i^Y). \quad (2)$$

The next step is to relate the density (2) to some mathematical object one is familiar with in statistical mechanics of disordered systems. This is done by means of the Edwards-Jones trick [9], which we briefly explain here. One first recall the following identity from distribution theory (or theory of generalised functions)

$$\lim_{\epsilon \rightarrow 0^+} \frac{1}{x - i\epsilon} = \text{P} \frac{1}{x} + i\pi\delta(x),$$

where  $\delta(x)$  is the Dirac delta. The expression above must be understood as distributions (that is, in the strictest mathematical way, under the integration sign and being multiplied by a test function). Using also that  $1/x = d \ln(x)/dx$  we can rewrite (2) as follows

$$\begin{aligned} \rho_Y(\lambda) &= \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi N(1)} \text{Im} \frac{\partial}{\partial z} \sum_{i=1}^{N(1)} \ln(z - \lambda_i^Y) \Big|_{z=\lambda-i\epsilon} \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi N(1)} \text{Im} \frac{\partial}{\partial z} \ln \prod_{i=1}^{N(1)} (z - \lambda_i^Y) \Big|_{z=\lambda-i\epsilon} \\ &= - \lim_{\epsilon \rightarrow 0^+} \frac{2}{\pi N(1)} \text{Im} \frac{\partial}{\partial z} \ln \left( \frac{1}{\sqrt{\det(z\mathbf{1}_{N(1)} - Y)}} \right) \Big|_{z=\lambda-i\epsilon} \end{aligned}$$

where  $\mathbf{1}_{N(1)}$  is the  $N(1) \times N(1)$  identity matrix. In this derivation we have used the algebraic result  $\prod_{i=1}^n \lambda_i^B = \det(B)$ , with  $\{\lambda_i^B\}_{i=1, \dots, n}$  being the eigenvalues of a  $n \times n$  matrix  $B$ . Notice also that we have rewritten  $\ln \det(z\mathbf{1}_{N(1)} - Y) = -2 \ln (1/\sqrt{\det(z\mathbf{1}_{N(1)} - Y)})$ . Since we know that for a positive definite matrix  $B$ , the expression  $1/\sqrt{\det(B)}$  can be expressed as a multi-dimensional Gaussian integral<sup>§</sup>, we

<sup>§</sup> It is also possible to stick with the expression  $\ln \det(z\mathbf{1}_{N(1)} - Y)$  and use Grassmann variables to write down an integral expression for  $\det(z\mathbf{1}_{N(1)} - Y)$ .

end up having the following expression for the density in (2):

$$\rho_Y(\lambda) = - \lim_{\epsilon \rightarrow 0^+} \frac{2}{\pi N(1)} \text{Im} \frac{\partial}{\partial z} \ln Z(z) \Big|_{z=\lambda-i\epsilon}, \quad (3)$$

$$\begin{aligned} Z(z) &= \frac{1}{\sqrt{\det(z\mathbf{1}_{N(1)} - Y)}} \\ &= \int \left[ \prod_{i=1}^{N(1)} \frac{dw_i(1)}{\sqrt{2\pi}} \right] \exp \left[ -\frac{1}{2} \sum_{i,j=1}^{N(1)} w_i(1)(z\mathbf{1}_{N(1)} - Y)_{ij} w_j(1) \right]. \end{aligned} \quad (4)$$

Expressions (3) and (4) entice us to understand  $\rho_Y(\lambda)$  as an observable corresponding to a system of  $N(1)$  continuous spins  $w_i(1)$  for  $i = 1, \dots, N(1)$  interacting with the Hamiltonian

$$H[\mathbf{w}(1)] = \frac{1}{2} \sum_{i,j=1}^{N(1)} w_i(1)(z\mathbf{1}_{N(1)} - Y)_{ij} w_j(1), \quad (5)$$

having a quenched random interaction matrix  $z\mathbf{1}_{N(1)} - Y$  and unit temperature. Of course one should not forget that  $z$  is generally complex, which strictly speaking implies that we do not have a Boltzmann measure. Yet, from a mathematical point of view, most (if not all) of the mathematical tricks that one can use in an interacting system defined by expressions (3) and (4) still apply, so one refers formally to  $Z(z)$  as the partition function (and the use of the jargon of *thermal averaging*, *Boltzmann measure*, et cetera swiftly follows).

Notice finally, that within this new context of interacting particles, the spectral density  $\rho_Y(\lambda)$  is related to the thermal average of the second moment of variable  $w_i(1)$ . Indeed, from results (3) and (4) we see that

$$\rho_Y(\lambda) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi N(1)} \text{Im} \sum_{i=1}^{N(1)} \langle w_i^2(1) \rangle_{z=\lambda-i\epsilon}, \quad (6)$$

$$\langle \cdots \rangle_z = \frac{1}{Z(z)} \int \left[ \prod_{i=1}^{N(1)} \frac{dw_i(1)}{\sqrt{2\pi}} \right] (\cdots) \exp \left[ -\frac{1}{2} \sum_{i,j=1}^{N(1)} w_i(1)(z\mathbf{1}_{N(1)} - Y)_{ij} w_j(1) \right]. \quad (7)$$

So far this approach is valid for any symmetric matrix  $Y$ .

We now turn our attention for  $Y$  matrices with the structure as given in (1). Here it is important to notice that for matrices of this sort, its entries are correlated. It is possible to treat the problem as it is, but it is much easier to disentangle the correlation of the entries of  $Y$  and express the problem in terms of the set of matrices  $X(t)$ . This is done by augmenting the system by adding new particles. The end result is the new thermal average  $\parallel$ :

$$\langle \cdots \rangle_z = \frac{1}{Z(z)} \int \left[ \prod_{t=1}^{M+1} d\mathbf{w}(t) \right] (\cdots) e^{-H[\mathbf{w}(1), \mathbf{w}(M+1)]} \prod_{t=1}^M W_t[\mathbf{w}(t+1)|\mathbf{w}(t)], \quad (8)$$

$\parallel$  Here, we ignore the  $2\pi$  factors as they are not important in the evaluation of the spectral density.

$$Z(z) = \int \left[ \prod_{t=1}^{M+1} d\mathbf{w}(t) \right] (\cdots) e^{-H[\mathbf{w}(1), \mathbf{w}(M+1)]} \prod_{t=1}^M W_t[\mathbf{w}(t+1)|\mathbf{w}(t)], \quad (9)$$

where we have defined

$$W_t[\mathbf{w}(t+1)|\mathbf{w}(t)] = \delta[\mathbf{w}(t+1) - X^T(t)\mathbf{w}(t)], \quad (10)$$

$$H[\mathbf{w}(1), \mathbf{w}(M+1)] = \frac{z}{2} \mathbf{w}^T(1)\mathbf{w}(1) - \frac{1}{2} \mathbf{w}^T(M+1)\mathbf{w}(M+1), \quad (11)$$

with  $\mathbf{w}(t) = (w_1(t), \dots, w_{N(t)}(t))$ , and where  $\delta(\mathbf{x}) = \prod_{i=1}^n \delta(x_i)$  for  $\mathbf{x} = (x_1, \dots, x_n)$ .

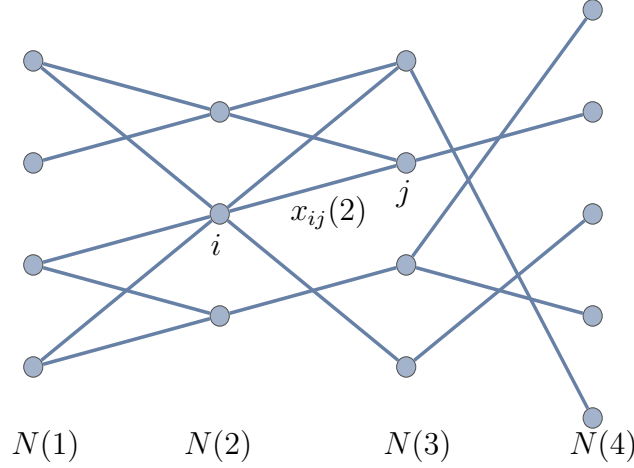
The new partition function (9) (and corresponding Boltzmann measure) is completely general for any set of rectangular matrices  $X(t)$ . However, due precisely to their rectangular character, finding a way to capture mathematically the most general ensemble embracing all possible scenarios seems a doomed task (e.g. how to properly define correlations between two different rectangular matrices  $X(t)$  and  $X(t')$ , or how to define symmetric properties of a given rectangular matrix  $X(t)$ ). Thus, we consider the simpler ensemble characterised by a lack of correlation between any distinct pair of matrices  $X(t)$  and  $X(t')$ . Besides, for each matrix  $X(t)$  its non-zero entries  $x_{ij}(t)$  are independent and identically distributed random variables with zero mean and variance  $J^2(t)$ . It is worth to point out that simple cases with correlation between matrices have been already studied [10]. The extension of the present work to consider a non-zero correlation matrix  $\Xi(t, t') = \langle X(t)X(t') \rangle$  is currently under way [11].

### 3. Cavity method and dense limit

In part I of this report we are going to use the cavity method<sup>¶</sup> to tackle the problem. For those interested in its replica counterpart, it will be used for completeness in part II. Note also that we are interested here in the properties of such matrices in the thermodynamic limit, that is, when the size of the matrices tends to infinity. Finite size corrections to the spectral density are possible by using e.g. the methodology of [16], but we postpone this analysis for another occasion.

Let us start by fixing some of the notation and giving some pictorial representation of the problem. We first notice that each matrix  $X(t)$  can be understood as a weighted bipartite graph in which a weighted link  $x_{ij}(t)$  connects a node  $i$  with  $i = 1, \dots, N(t)$  with a node  $j$  with  $j = 1, \dots, N(t+1)$ . For a node on the  $t$ -layer (or a  $t$ -node)  $x_{ij}(t)$  can be understood as an outgoing link while this will be considered as an incoming link for a  $(t+1)$ -node. As we have  $M$  matrices  $X(t)$ , the overall graph is a combination of coupled bipartite graphs between consecutive  $t$ -layers (also called a multipartite graph), as depicted in figure 1. Now, for each matrix  $X(t)$  the total number of outgoing links in the  $t$ -layer must match the number of incoming links in the  $(t+1)$ -layer. Let us denote as  $k_{i,t}$  the number of outgoing links from node  $i$  in the  $t$ -layer, and as

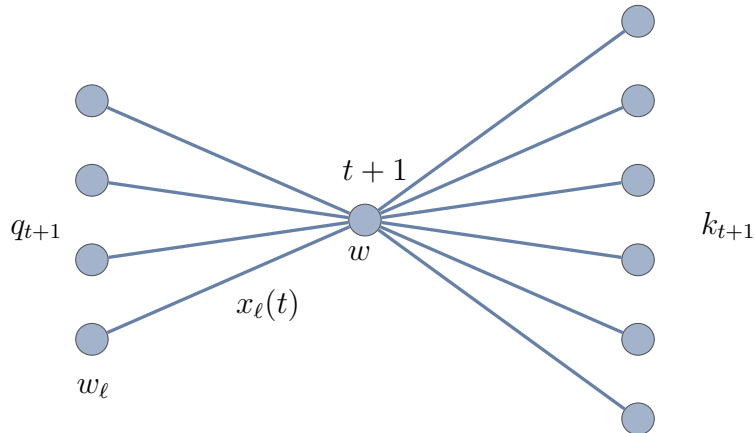
<sup>¶</sup> See, for instance, [12] for generalities of the cavity method or [13, 14] for simpler applications of the cavity method in similar problems. Another useful reference to understand the cavity method is [15]



**Figure 1.** Pictorial representation of a multipartite graph with  $N(1) = 4$ ,  $N(2) = 3$ ,  $N(3) = 4$  and  $N(4) = 5$

$q_{j,t+1}$  the number of incoming links to node  $j$  in the  $(t+1)$ -layer. We must have that  $\sum_{i=1}^{N(t)} k_{i,t} = \sum_{j=1}^{N(t+1)} q_{j,t+1}$ . Alternatively we write  $N(t)\bar{k}_t = N(t+1)\bar{q}_{t+1}$ , with  $\bar{k}_t = (1/N(t)) \sum_{i=1}^{N(t)} k_{i,t}$  and  $\bar{q}_{t+1} = (1/N(t+1)) \sum_{j=1}^{N(t+1)} q_{j,t+1}$  the average outgoing in incoming connectivities, respectively.

Consider next how the situation looks like around one of the  $(t+1)$ -nodes (see figure 2). For simplicity we drop the indices  $i, j$  and work with generic nodes. Here a generic node in the  $(t+1)$ -layer with variable  $w$  has  $q_{t+1}$  incoming links with weights  $\{x_\ell(t)\}_{\ell=1}^{q_{t+1}}$  and  $k_{t+1}$  outgoing links. To apply the cavity method we focus in the generic case



**Figure 2.** Pictorial representation of a local node with variable  $w$  at the  $(t+1)$ -layer in a multipartite graph

depicted in figure 2 and try to derive marginal distributions evaluating the constraint  $\delta(w - \sum_{\ell=1}^{q_{t+1}} x_\ell(t)w_\ell)$ . This constraint states that  $q_{t+1}$  variables  $\{w_\ell\}_{\ell=1}^{q_{t+1}}$  at the  $t$ -layer will determine the value of  $w$  of the  $(t+1)$ -node. This value will in turn participate in

$k_{t+1}$  other constraints.

Suppose we fix the value of  $w$  at the  $(t+1)$ -node and wonder about its impact to the left and to the right. Let us denote as  $P_t$  the marginals at the  $t$ -layer resulting from this operation, while the marginals to the right will be denoted as  $Q_{t+1}$ . These marginals are obviously different as the role that the variable  $w$  at the  $(t+1)$ -node plays to the left and to the right is different. It is possible to write down recursive relations for them. Indeed, for the  $Q$ -marginals we have that

$$Q_t(w) = \frac{1}{Q_t} \left[ \int dw' \prod_{\ell=1}^{k_{t+1}} Q_{t+1,\ell}(w') \right] \left[ \int \prod_{\ell=1}^{q_{t+1}-1} dw_\ell P_{t,\ell}(w_\ell) \right] \times \delta \left( w' - x(t)w - \sum_{\ell=1}^{q_{t+1}-1} x_\ell(t)w_\ell \right), \quad (12)$$

for  $t = 1, \dots, M-1$ . Similarly, for the  $P$ -marginals one finds

$$P_{t+1}(w) = \frac{1}{P_{t+1}} \left[ \prod_{\ell=1}^{k_{t+1}-1} Q_{t+1,\ell}(w) \right] \left[ \int \prod_{\ell=1}^{q_{t+1}} dw_\ell P_{t,\ell}(w_\ell) \right] \delta \left( w - \sum_{\ell=1}^{q_{t+1}} x_\ell(t)w_\ell \right), \quad (13)$$

for  $t = 1, \dots, M-1$ . These equations do not apply to the first and last layers as their neighbourhood is rather different: nodes in the first layer only have outgoing links, while nodes in the last layer only have incoming links. Moreover, the nodes in the last layer carry a weight  $e^{w^2/2}$ , as indicated by the partition function. Thus in this case the corresponding equations are:

$$P_1(w) = \frac{e^{-\frac{z}{2}w^2}}{P_1} \prod_{\ell=1}^{k_1-1} Q_{1,\ell}(w) \\ Q_M(w) = \frac{1}{Q_M} \left[ \int dw' e^{\frac{1}{2}(w')^2} \right] \left[ \int \prod_{\ell=1}^{q_{M+1}-1} dw_\ell P_{M,\ell}(w_\ell) \right] \times \delta \left( w' - x(M+1)w - \sum_{\ell=1}^{q_{M+1}-1} x_\ell(M+1)w_\ell \right) \quad (14)$$

The set of equations (12), (13), and (15) are our cavity equations for the  $Q$  and  $P$  marginals. Once they are solved, any statistical quantity of interest can be estimated from them. In particular, the spectral density is related to the second moment of the actual marginals in the first layer. These are given by

$$\mathcal{P}(w) = \frac{e^{-\frac{z}{2}w^2}}{\mathcal{P}} \prod_{\ell=1}^{k_1} Q_{1,\ell}(w)$$

As it has been noticed in previous works [13, 14, 17], the subset of Gaussian distributions is a functional fixed-point of the cavity equations. Thus, parametrising the cavity marginals as  $Q_t(w) = \frac{1}{\sqrt{2\pi/(x^2(t)\gamma(t))}} e^{-\frac{x^2(t)\gamma(t)}{2}w^2}$  and  $P_t(w) = \frac{1}{\sqrt{2\pi\sigma(t)}} e^{-\frac{w^2}{2\sigma(t)}}$  permits us to transform the original functional equations into standard equations for variances;

$$\gamma(t) = \left( \sum_{\ell=1}^{q_{t+1}-1} x_\ell^2(t)\sigma_\ell(t) + \frac{1}{\sum_{\ell=1}^{k_{t+1}} x_\ell^2(t+1)\gamma_\ell(t+1)} \right)^{-1}, \quad t = 1, \dots, M-1$$

$$\begin{aligned}
\gamma(M) &= -\frac{1}{1 - \sum_{\ell=1}^{q_{M+1}-1} x_\ell^2(M) \sigma_\ell(M)}, \\
\sigma(1) &= \frac{1}{z + \sum_{\ell=1}^{k_1-1} x_\ell^2(1) \gamma_\ell(1)}, \\
\sigma(t+1) &= \frac{1}{\sum_{\ell=1}^{k_{t+1}-1} x_\ell^2(t+1) \gamma_\ell(t+1) + \frac{1}{\sum_{\ell=1}^{q_{t+1}-1} x_\ell^2(t) \sigma_\ell(t)}}, \quad t = 1, \dots, M-1,
\end{aligned} \tag{15}$$

while for the actual variance, the one parametrising  $\mathcal{P}$ , we obtain  $\Delta = [z + \sum_{\ell=1}^{k_1} x_\ell^2(1) \gamma_\ell(1)]^{-1}$ . We keep the set of equations (15) for a generic node as they are, even though they can easily be rewritten for a given instance of a multipartite graph. These equations are of course valid for typical and large multipartite graph where short loops are rare, and can be solved numerically by using fixed point iteration methods (also called belief-propagation in this context) to render the spectral density of any large instance. Alternatively, one may prefer to use the ensemble equations and used population dynamics instead. This, and exact solutions for random regular multipartite graphs, will be the topic of part II of this work.

Here we focus on the dense case that, as we will see, it is already quite challenging to deal with. To perform the dense limit  $k_t, q_t \rightarrow \infty$  in (15), we first rescale the entries  $x_\ell(t) \rightarrow x_\ell(t)/\sqrt{q_{t+1}}$ . In the subsequent derivations we encounter expressions of the sort  $\frac{1}{q_{t+1}} \sum_{\ell=1}^{q_{t+1}-1} x_\ell^2(t) \sigma_\ell(t) \rightarrow J^2(t) \sigma(t)$ . Here several steps have been done at the same time in the limit for large  $q_{t+1}$  and  $k_t$ , the important ones being: (i) the variables  $x_\ell^2(t)$  and  $\sigma_\ell(t)$  decorrelate; (ii) no Onsager's reaction term is generated; (iii) nodes are equivalent within each layer  $t$ <sup>+</sup>. Taking this into account, we obtain the following set of equations in the dense (sometimes rightfully called fully connected) limit:

$$\begin{aligned}
\Sigma(t) &= \frac{\alpha_t}{\alpha_{t-1}} \left( \frac{1}{\sigma(t)} - \frac{1}{J^2(t)} \frac{1}{\sigma(t)} \sigma(t+1) \frac{1}{\sigma(t)} \right), \quad t = 1, \dots, M-1, \\
\Sigma(M) &= -\frac{\alpha_M}{\alpha_{M-1}} \frac{J^2(M)}{1 - J^2(M) \sigma(M)} \\
\sigma(1) &= \frac{1}{z + \Sigma(1)}, \quad \sigma(t+1) = \frac{1}{\Sigma(t+1) + \frac{1}{J^2(t) \sigma(t)}}, \quad t = 1, \dots, M-1,
\end{aligned}$$

where we have used that  $k_t = q_{t+1} \frac{\alpha_t}{\alpha_{t-1}}$  with  $\alpha_{t-1} = N(t)/N(1)$  -notice that  $\alpha_0 = 1$ - and defined  $\Sigma(t) \equiv \frac{\alpha_t}{\alpha_{t-1}} J^2(t) \gamma(t)$ .

It is possible, as it was noticed in [2] by other means, to show that  $\sigma(1)$  must obey the polynomial equation (see its derivation in appendix Appendix A):

$$P_{M+1}(v) = vz, \quad P_{M+1}(v) = \prod_{s=1}^{M+1} \left( 1 + \frac{\alpha_M}{\alpha_{s-1}} v \right), \quad v = -\frac{1 - z\sigma(1)}{\alpha_M} \tag{16}$$

<sup>+</sup> Onsager's reaction term is the correcting term that appears when determining the value of a random variable  $x_i$ , whose value depends e.g. on a sum of random variables in which the variable  $x_i$  is also involved. If the typical value of the sum is zero, any change in  $x_i$  will determine the sign of the sum and therefore one must account for changes in  $x_i$  (thus Onsager's reaction term). If, however, the typical value of the sum is different from zero, changes in  $x_i$  will not affect the sum. In this case, there is not Onsager's reaction term, which is precisely our situation here. Further intuition behind this effect is nicely seen in, for instance, [15] in models of disordered magnets.



where we have rescaled  $z/\mathcal{J}_M^2 \rightarrow z$  and  $\mathcal{J}_M^2\sigma(1) \rightarrow \sigma(1)$ , with  $\mathcal{J}_M = \prod_{s=1}^M J(s)$ . Finally, and since in the dense limit  $\Delta = \sigma(1)$ , the spectral density is obtained using the formula:

$$\rho(\lambda) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \text{Im} [\sigma(1)]|_{z=\lambda-i\epsilon} \quad (17)$$

The rest of this work consists in analysing the solutions coming from the polynomial equation (16). Some words are in order, though. It is evidently clear that solutions by radicals can be found for  $M$  up to three, but even for  $M = 3$  the resulting expressions are rather unpleasant. It is however possible to consider solutions for a particular set of parameters for which the resulting expressions are much better-looking, while still keeping their main properties somewhat intact. With this in mind, let us introduce some notation. Recalling that  $X(1) \cdots X(M)$ , we denote as  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_M)$  the standard choice of parameters and as  $\boldsymbol{\alpha}^{(s)} = (1, \overset{s-1}{\cdot}, 1, \alpha, \dots, \alpha)$  the choice of parameters to describe the situation in which the first  $s-1$ -matrices are square of unit size, the  $s$ -matrix is rectangular of relative size  $1 \times \alpha$  and the rest are square of size  $\alpha \times \alpha$ , that is:

$$1 \left\{ \overbrace{\begin{pmatrix} \cdots \\ \vdots & \ddots & \vdots \\ \cdots \end{pmatrix}}^1 \times \overbrace{\begin{pmatrix} \cdots \\ \vdots & \ddots & \vdots \\ \cdots \end{pmatrix}}^1 \times \cdots \times \underbrace{\overbrace{\begin{pmatrix} \cdots \\ \vdots & \ddots & \vdots \\ \cdots \end{pmatrix}}^{\alpha}}_{s\text{-matrix}} \times \overbrace{\begin{pmatrix} \cdots \\ \vdots & \ddots & \vdots \\ \cdots \end{pmatrix}}^{\alpha} \times \cdots \times \overbrace{\begin{pmatrix} \cdots \\ \vdots & \ddots & \vdots \\ \cdots \end{pmatrix}}^{\alpha} \right.$$

#### 4. Exact results by radicals for $M = 1, 2$ and $3$

##### 4.1. Case $M = 1$

For  $M = 1$ , after solving the quadratic equation for  $\sigma(1)$  and extracting from it the spectral density using (17), we obtain the following expression:

$$\rho(\lambda) = (1 - \min(1, \alpha_1))\delta(\lambda) + \frac{1}{2\pi\lambda} \sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)} \mathbf{1}_{\lambda \in [\lambda_-, \lambda_+]}$$

with  $\lambda_{\pm} = (1 \pm \sqrt{\alpha_1})^2$ . This is as expected the Marcenko-Pastur law. We would like to point out, as this will be used later on, that for small  $\alpha_1$ , the continuous part of the spectrum tends to a Wigner law after proper scaling. Indeed, for a small  $\alpha_1$  we have that  $\lambda_{\pm} = 1 \pm 2\sqrt{\alpha_1} + \dots$ . If we define  $\lambda = 1 + 2\sqrt{\alpha_1}w$  we have that the continuous part gives goes to  $\alpha_1 \frac{2}{\pi} \sqrt{1 - w^2}$  for  $w \in [-1, 1]$ .

##### 4.2. Case $M = 2$

After analysing the cubic polynomial in detailed and looking for the best way to simplify the expressions (see appendix Appendix B) we find that the spectral density reads:

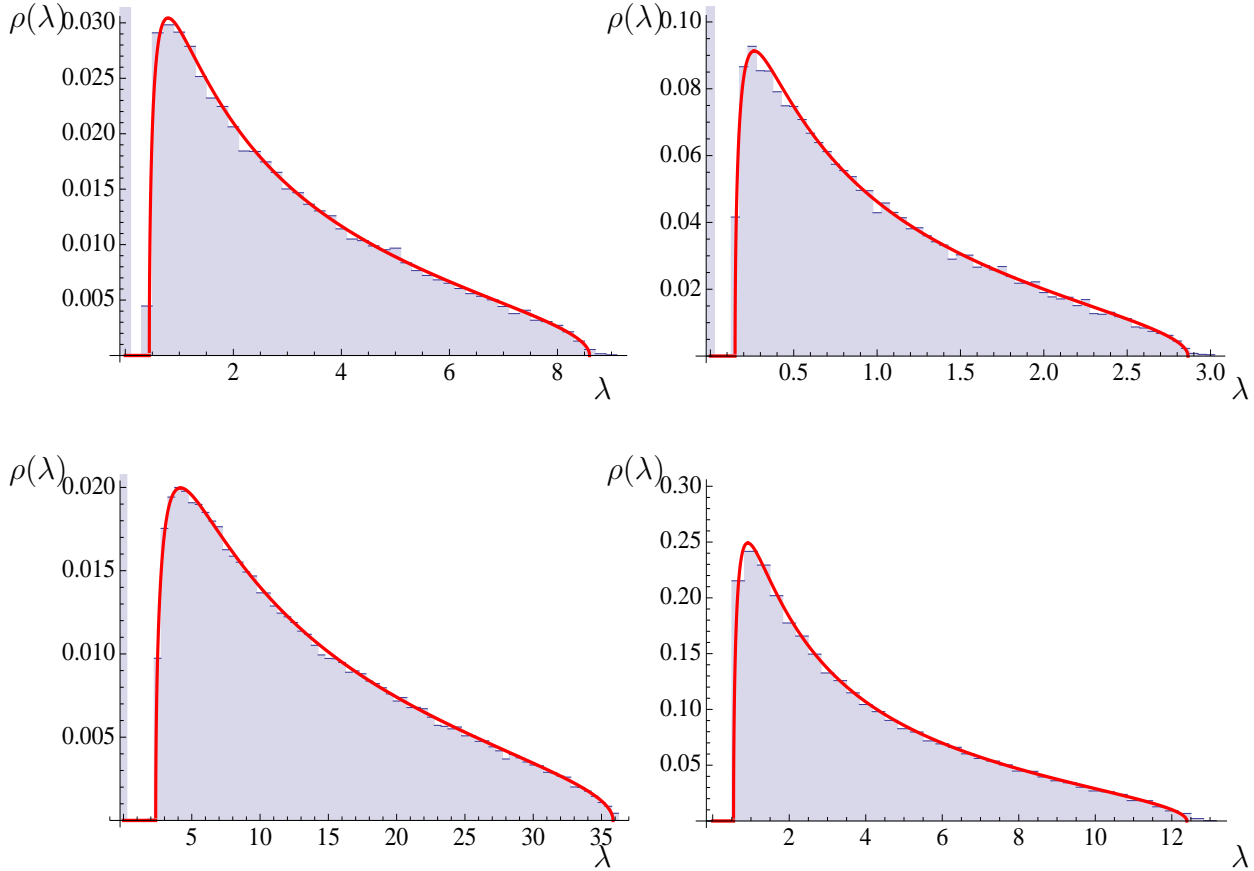
$$\begin{aligned} \rho(\lambda) = & (1 - \min(1, \alpha_1, \alpha_2))\delta(\lambda) \\ & + \frac{\sqrt{3}}{6\pi\lambda\sqrt[3]{2}} \left( \sqrt[3]{9\alpha_1\gamma(\lambda - \xi_0) + 6\sqrt{3\alpha_1^3(\lambda - \lambda_-)(\lambda_{+,2} - \lambda)(\lambda - \lambda_{+,1})}} \right. \\ & \left. - \sqrt[3]{9\alpha_1\gamma(\lambda - \xi_0) - 6\sqrt{3\alpha_1^3(\lambda - \lambda_-)(\lambda_{+,2} - \lambda)(\lambda - \lambda_{+,1})}} \right) \mathbf{1}_{\lambda \in [\lambda_{+,1}, \lambda_{+,2}]} \end{aligned}$$

with

$$\xi_0 = -\frac{2(1 + \alpha_1 + \alpha_2)^3 - 9(\alpha_1(1 + \alpha_1) + \alpha_2(1 + \alpha_2) + \alpha_1\alpha_2(\alpha_1 + \alpha_2))}{9\alpha_1(1 + \alpha_1 + \alpha_2)}$$

Here  $\lambda_-$ ,  $\lambda_{+,1}$ , and  $\lambda_{+,2}$  are the roots of a cubic the cubic polynomial reported in the appendix Appendix B. We do not report their explicit form for the general case, as it suffices to point out that: (i) since  $(1 + \alpha_1 + \alpha_2)^3 > 27\alpha_1\alpha_2$ , the three roots are real; (ii) for the cases  $\alpha_1 = 1$  or  $\alpha_2 = 1$  or  $\alpha_1 = \alpha_2$ , the polynomial coefficient  $d = 0$ , meaning that one root is zero. Any of these choices correspond precisely to the family of parameters mentioned above; (iii) generally, there is one a non-positive root  $\lambda_-$  and two non-negatives roots  $\lambda_{+,1}$  and  $\lambda_{+,2}$  with  $\lambda_{+,2} > \lambda_{+,1}$ .

In figure 3 we have plotted the spectral density for four different values of the pair of parameters  $(\alpha_1, \alpha_2)$  and compared them with numerical diagonalisation.



**Figure 3.** Clockwise with notation  $(\alpha_1, \alpha_2)$  we have:  $(0.1, 0.3)$  (for histogram we have  $N(1) = 1000$ ,  $N(2) = 100$  and  $N(3) = 300$  averaged over 1000 samples),  $(0.3, 0.1)$  (for histogram we have  $N(1) = 1000$ ,  $N(2) = 300$  and  $N(3) = 100$  averaged over 1000 samples),  $(4, 4)$  (with  $N(1) = 100$ ,  $N(2) = 400$  and  $N(3) = 400$  averaged over 1000 samples), and  $(0.3, 4)$  ( $N(1) = 500$ ,  $N(2) = 150$  and  $N(3) = 2000$  averaged 1000 samples).

From the analysis of the roots and the expression of the continuous part of the spectral density we have that  $\rho_{cont}(w, \alpha_1, \alpha_2) = (\alpha_1/\alpha_2)\rho_{cont}((\alpha_1/\alpha_2)w, \alpha_2, \alpha_1)$ , scaling

that, by simple eye inspection, can be numerically observed in figure 3.

Let us move on to analyse the cases for which the polynomial coefficient  $d = 0$ .

**4.2.1.  $\alpha^{(1)} = (\alpha, \alpha)$  (first rectangular, second square)** This case is interesting as it yields a different scenario to the one observed for the Marcenko-Pastur law, namely, having no gap and Dirac delta contribution at zero. Indeed, the spectral density here reads:

$$\begin{aligned} \rho(\lambda) &= (1 - \min(1, \alpha)) \delta(\lambda) \\ &+ \frac{\sqrt{3}}{6\pi\lambda\sqrt[3]{2}} \left( \sqrt[3]{9\alpha(1+2\alpha)(\lambda - \xi_0) + 6\sqrt{3\alpha^3(\lambda - \lambda_-)\lambda(\lambda_+ - \lambda)}} \right. \\ &\quad \left. - \sqrt[3]{9\alpha(1+2\alpha)(\lambda - \xi_0) - 6\sqrt{3\alpha^3(\lambda - \lambda_-)\lambda(\lambda_+ - \lambda)}} \right) \mathbf{1}_{\lambda \in [\lambda_-, \Theta(\alpha-1), \lambda_+]} \end{aligned}$$

with

$$\lambda_{\pm} = \frac{-1 + 20\alpha + 8\alpha^2 \pm (1 + 8\alpha)^{3/2}}{8\alpha}, \quad \xi_0 = \frac{2(-1 + \alpha)^3}{9\alpha(1 + 2\alpha)}.$$

Notice that  $\lambda_- \leq 0$  for  $\alpha \in (0, 1]$ , while  $\lambda_-$  is positive for  $\alpha \in (1, \infty)$ . At the same time  $\lambda_+ \geq 4$  from  $\alpha \in [0, \infty)$ . Besides  $\lambda_- \simeq -\frac{1}{4\alpha}$  as  $\alpha \rightarrow 0$ . The attentive reader surely have realised that for  $\alpha \in (0, 1)$  there is both a continuous density at zero as well as a Dirac delta contribution, while for  $\alpha > 1$  the Dirac delta disappears and a gap appears in the continuous part (see figure 4).

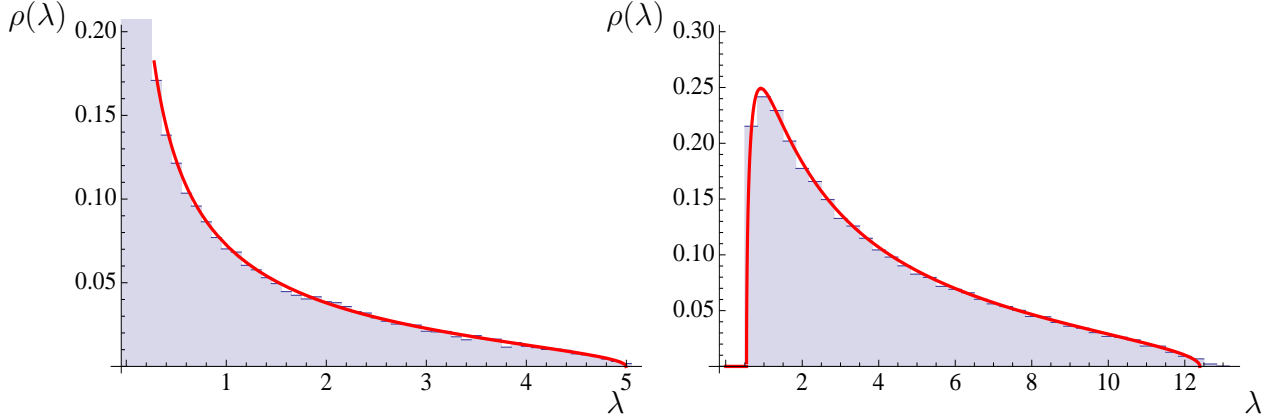
For small  $\alpha$  we can see better the combined contribution at zero of the continuous and isolated part. Moreover, the continuous part of the spectrum disappears with a factor  $\alpha$  while keeping a well-defined shape reminiscent of the Marcenko-Pastur law for  $\alpha = 1$ , viz.

$$\rho(\lambda) = (1 - \alpha) \delta(\lambda) + \frac{\alpha}{2\pi} \sqrt{\frac{4 - \lambda}{\lambda}} \mathbf{1}_{\lambda \in [0, 4]} + \mathcal{O}(\alpha^2)$$

The curious reader would have noticed that the first term in  $\alpha$  is such that the spectral density is normalised to one. One wonders what happens to the normalisation of higher-order terms in the expansion: first of all the functions of higher-order terms are not densities (they can generally have positive and negative values); secondly, apart of the  $\alpha^2$  correction whose integral is zero, the integral of the rest of the terms diverges term by term.

**4.2.2.  $\alpha^{(2)} = (1, \alpha)$  (first matrix square, second rectangular)** In this case, the spectral density becomes:

$$\begin{aligned} \rho(\lambda) &= (1 - \min(1, \alpha)) \delta(\lambda) \\ &+ \frac{\sqrt{3}}{6\pi\lambda\sqrt[3]{2}} \left( \sqrt[3]{9(2 + \alpha)(\lambda - \xi_0) + 6\sqrt{3(\lambda - \lambda_-)\lambda(\lambda_+ - \lambda)}} \right. \\ &\quad \left. - \sqrt[3]{9(2 + \alpha)(\lambda - \xi_0) - 6\sqrt{3(\lambda - \lambda_-)\lambda(\lambda_+ - \lambda)}} \right) \mathbf{1}_{\lambda \in [\lambda_-, \Theta(1-\alpha), \lambda_+]} \end{aligned}$$

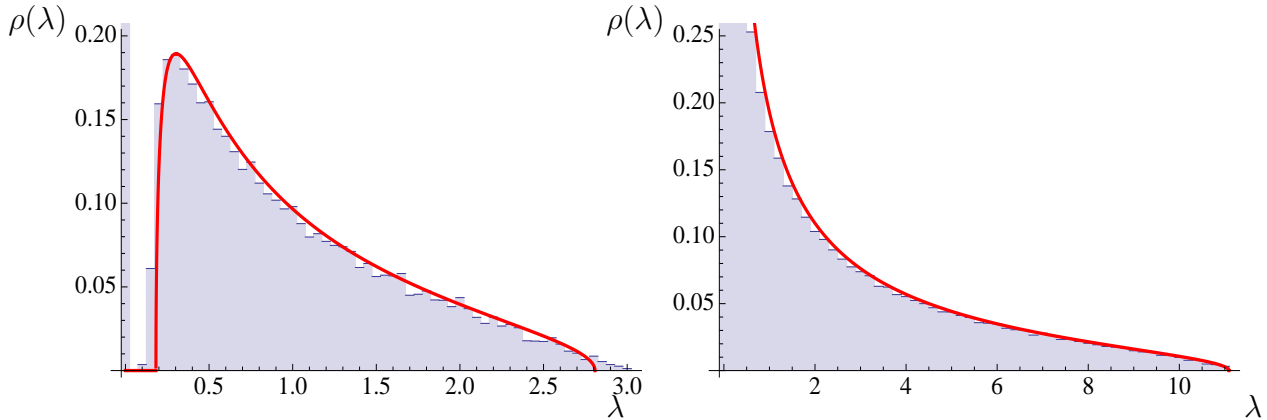


**Figure 4.** Results for  $\alpha^{(1)} = (\alpha, \alpha)$  with  $\alpha = 3/10$  (histograms have been performed for  $N(1) = 500$ ,  $N(2) = N(3) = 150$  averaged over 1000 samples) at the left where there is also a Dirac delta contribution at zero and for  $\alpha = 4$  (histograms have been performed for  $N(1) = 100$ ,  $N(2) = N(3) = 400$  averaged over 1000 samples) at the right plot.

with

$$\lambda_{\pm} = \frac{1}{8}(8 + 20\alpha - \alpha^2 \pm \sqrt{\alpha}(8 + \alpha)^{3/2}), \quad \xi_0 = -\frac{2(-1 + \alpha)^3}{9(2 + \alpha)}$$

We have that  $\lambda_+ \geq \lambda_-$  the equality only resulting when  $\alpha = 0$  in which case  $\lambda_- = \lambda_+ = 1$ . Now, we have that  $\lambda_+$  is always positive while  $\lambda_-$  is positive for  $\alpha \leq 1$  and then  $\lambda_- < 0$  for  $\alpha \in (1, \infty)$ . This implies that there will be a gap for  $\alpha \in [0, 1]$  with Dirac delta, while for  $\alpha > 1$  there is no gap and no Dirac delta, which again differs from the well-known properties of the Marcenko-Pastur law. In the figure 5 we plot the two cases.



**Figure 5.** Case  $\alpha^{(2)} = (1, \alpha)$  for two values of  $\alpha = 2/10$  (for the histograms we took  $N(1) = 100$ ,  $N(2) = 100$  and  $N(3) = 20$  averaged over 1000 samples) and  $\alpha = 2$  (for the histograms we took  $N(1) = 100$ ,  $N(2) = 100$  and  $N(3) = 200$  averaged over 1000 samples)

It is interesting to see what happens for small  $\alpha$ , since this asymptotic behaviour, as we will see later on, depends strongly on the position of the rectangular matrix. We note that for small  $\alpha$  the endpoints have the expansion  $\lambda_- = 1 - 2\sqrt{2}\sqrt{\alpha} + \dots$ ,  $\lambda_+ = 1 + 2\sqrt{2}\sqrt{\alpha} + \dots$ . This implies that the continuous part of the spectrum shrinks to zero changing its shape, unlike the previous case for  $\alpha^{(1)}$ . To obtain the limiting vanishing shape, we do the change of variables  $\lambda = 2\sqrt{2\alpha}w + 1$  to obtain that  $\rho_{cont} \rightarrow \alpha \frac{2}{\pi} \sqrt{1-w^2}$ , that is, the Wigner law.

Let us finish by noticing that, similarly to the symmetry of the Marcenko-Pastur law relating  $\alpha \leftrightarrow 1/\alpha$ , we have here that the spectral densities corresponding to the two choices of parameters are related by a transformation  $(\lambda, \alpha) \rightarrow (\lambda/\alpha, 1/\alpha)$ .

#### 4.3. Case $M = 3$

Finally, for  $M = 3$  the spectral density is given by the following expression

$$\rho(\lambda) = (1 - \min(1, \alpha_1, \alpha_2, \alpha_3))\delta(\lambda) + \frac{1}{2\pi\lambda} \sqrt{4\mathcal{S}^2(\lambda) + \frac{\gamma^2 - 4\gamma_2}{4} + \frac{1}{8} \frac{Q_1(\lambda)}{\mathcal{S}(\lambda)}} \mathbf{1}_{\lambda \in [\lambda_+, 1, \lambda_+, 2]}$$

with

$$\mathcal{S}(\lambda) = \frac{1}{2} \sqrt{-\frac{\gamma^2 - 4\gamma_2}{12} + \frac{1}{3} \left( \sqrt[3]{\frac{P_2(\lambda) + \sqrt{P_4(\lambda)}}{2}} + \frac{P_1(\lambda)}{\sqrt[3]{\frac{P_2(\lambda) + \sqrt{P_4(\lambda)}}{2}}} \right)}$$

where the derivation and definitions of the polynomials involved can be found in appendix Appendix C. Here  $\lambda_-$  and  $\lambda_+$  are roots from solving  $4\mathcal{S}^2(\lambda) + \frac{\gamma^2 - 4\gamma_2}{4} + \frac{1}{8} \frac{Q_1(\lambda)}{\mathcal{S}(\lambda)} = 0$ . These two roots seems to be related to the zeros of polynomial  $P_4(\lambda)$ , albeit in a complicated manner at least for the lower endpoint.

As the general solution is fairly involved mathematically, we look for a choice of parameters which simplify the zeros of  $P_4$ . This choice turns out to be the ones we introduced above.

4.3.1.  $\alpha^{(1)} = (\alpha, \alpha, \alpha)$  (*recangular-square-square*) Particularising for this case we have the following spectral density

$$\rho(\lambda) = (1 - \min(1, \alpha))\delta(\lambda) + \frac{1}{2\pi\lambda} \sqrt{4\mathcal{S}^2(\lambda) - \frac{3(1-\alpha)^2}{4} - \frac{\alpha^2 \left( \lambda + \frac{(-1+\alpha)^3}{8\alpha^2} \right)}{\mathcal{S}(\lambda)}} \mathbf{1}_{\lambda \in [\Theta(\alpha-1)\lambda_-, \lambda_+]}$$

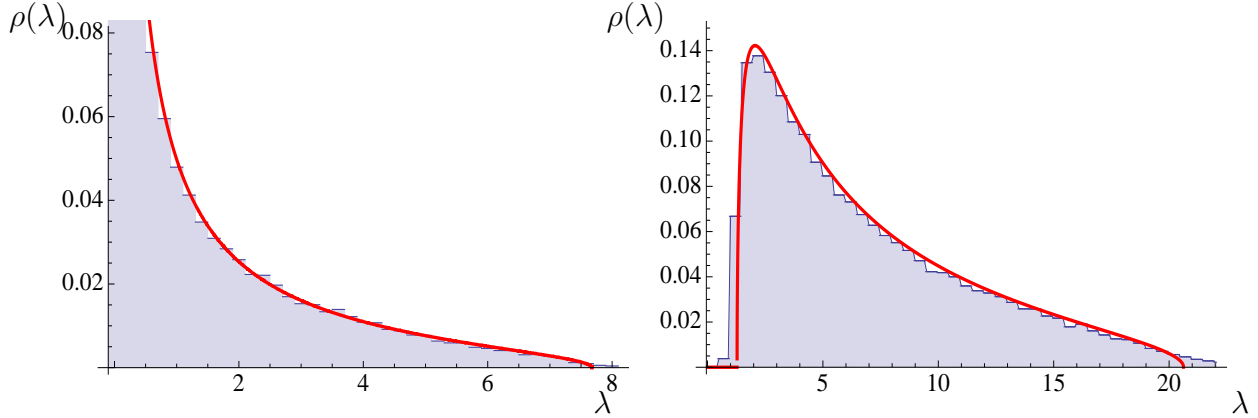
with

$$\mathcal{S}(\lambda) = \frac{1}{2} \left( \frac{(1-\alpha)^2}{4} + \left( \frac{\alpha^4 \lambda}{2} \right)^{1/3} \sqrt[3]{\lambda + \frac{(1-\alpha)^2}{\alpha} + \sqrt{(\lambda - \lambda_-)(\lambda - \lambda_+)}} + \frac{\alpha^{2/3}(1+3\alpha)\lambda}{3 \left( \frac{\lambda}{2} \right)^{1/3} \sqrt[3]{\lambda + \frac{(1-\alpha)^2}{\alpha} + \sqrt{(\lambda - \lambda_-)(\lambda - \lambda_+)}}} \right)^{1/2}$$

and

$$\lambda_{\pm} = \frac{2 \pm 2\sqrt{(1+3\alpha)^3(-1+9\alpha)^2} + 9\alpha(-1+3\alpha(4+\alpha))}{27\alpha^2}$$

In the figure 6 we show some plots and their comparison with numerical diagonalisation.



**Figure 6.** Results for  $\alpha^{(1)} = (\alpha, \alpha, \alpha)$ . Left:  $\alpha = 3/10$  and there is a Dirac delta not visible as it mixes up with the continuous part (for the histograms we have chosen  $N(1) = 500$  and  $N(2) = N(3) = N(4) = 150$  averaged over 1000 samples). Right:  $\alpha = 7$  (for the histograms we have chosen  $N(1) = 10$  and  $N(2) = N(3) = N(4) = 70$  averaged over 1000 samples).

The astute reader has surely already realised what is to be expected when considering small  $\alpha$ , by simply looking at the number of square matrices of size  $\alpha$ . Indeed for small  $\alpha$  we obtain

$$\begin{aligned} \rho(\lambda) &= (1 - \alpha)\delta(\lambda) \\ &+ \frac{\alpha}{2^{4/3}3^{1/6}\pi\lambda} \left[ \left(9\lambda + \sqrt{3}\lambda\sqrt{27-4\lambda}\right)^{1/3} - \left(9\lambda - \sqrt{3}\lambda\sqrt{27-4\lambda}\right)^{1/3} \right] \mathbf{1}_{\lambda \in [0, \frac{27}{4}]} \\ &+ \mathcal{O}(\alpha^2), \end{aligned}$$

that is, the continous part keeps a definite shape -the one corresponding to the spectral density of  $M = 2$  for  $\alpha_1 = \alpha_2 = 1$ - while vanishing with an overall factor  $\alpha$ .

#### 4.4. Case $\alpha^{(2)} = (1, \alpha, \alpha)$ (square-rectangular-square)

Particularising the formulas in appendix Appendix C for this case we obtain

$$\rho(\lambda) = (1 - \min(1, \alpha))\delta(\lambda) + \frac{1}{2\pi\lambda} \sqrt{4\mathcal{S}^2(\lambda) - (1 - \alpha)^2 - \frac{\alpha\lambda}{\mathcal{S}(\lambda)}} \mathbf{1}_{\lambda \in [0, \lambda_+]}$$

with

$$\mathcal{S}(\lambda) = \frac{1}{2\sqrt{3}} \left( (1 - \alpha)^2 + 3\alpha^{2/3} \sqrt[3]{\frac{(\lambda - p_-)(\lambda - p_+) + \lambda\sqrt{(\lambda - \lambda_-)(\lambda - \lambda_+)}}{2}} \right)$$

$$+ \frac{(1 - \alpha)^4 + 6\alpha(1 + \alpha)\lambda}{3\alpha^{2/3} \sqrt[3]{\frac{(\lambda - p_-)(\lambda - p_+) + \lambda\sqrt{(\lambda - \lambda_-)(\lambda - \lambda_+)}}{2}}})^{1/2}$$

and

$$\lambda_{\pm} = \frac{2 \left( \alpha(33 - (\alpha - 33)\alpha) \pm \sqrt{(\alpha(\alpha + 14) + 1)^3 - 1} \right)}{27\alpha},$$

$$p_{\pm} = \frac{-3(\alpha - 1)^2\alpha(\alpha + 1) \pm \sqrt{3}\sqrt{(\alpha - 1)^4\alpha^2(\alpha(\alpha + 10) + 1)}}{9\alpha^2}.$$

This case yields a situation where there is *no gap for any value of  $\alpha$* . Actually, as we will see later this *no gap property*, it is generally observed for the family of parameter  $\alpha^{(s)}$  for  $s = 2, \dots, M - 1$ .

For small  $\alpha$  we again see that the continuous part of the spectral density keeps a well-defined shape, viz

$$\rho(\lambda) = (1 - \alpha)\delta(\lambda) + \alpha \frac{1}{4\pi} \sqrt{\frac{4 - \lambda}{\lambda}} + \dots$$

which corresponds, as expected, to the spectral density for  $M = 1$  and  $\alpha_1 = 1$ .

*4.4.1. Case  $\alpha^{(3)} = (1, 1, \alpha)$  (square-square-rectangular)* For this final choice of parameter we obtain

$$\rho(\lambda) = (1 - \min(1, \alpha))\delta(\lambda) + \frac{1}{2\pi\lambda} \sqrt{4\mathcal{S}^2(\lambda) - \frac{3(1 - \alpha)^2}{4} - \frac{1}{8} \frac{8\lambda + (1 - \alpha)^3}{\mathcal{S}(\lambda)}} \mathbf{1}_{\lambda \in [\lambda_-, \Theta(1 - \alpha), \lambda_+]}$$
(18)

with

$$\mathcal{S}(\lambda) = \frac{1}{2} \left( \frac{(1 - \alpha)^2}{4} + \lambda^{1/3} \sqrt[3]{\frac{\lambda + (1 - \alpha)^2 + \sqrt{(\lambda - \lambda_-)(\lambda - \lambda_+)}}{2}} \right. \\ \left. + \frac{(3 + \alpha)\lambda^{2/3}}{3 \sqrt[3]{\frac{\lambda + (1 - \alpha)^2 + \sqrt{(\lambda - \lambda_-)(\lambda - \lambda_+)}}{2}}} \right)^{1/2}$$

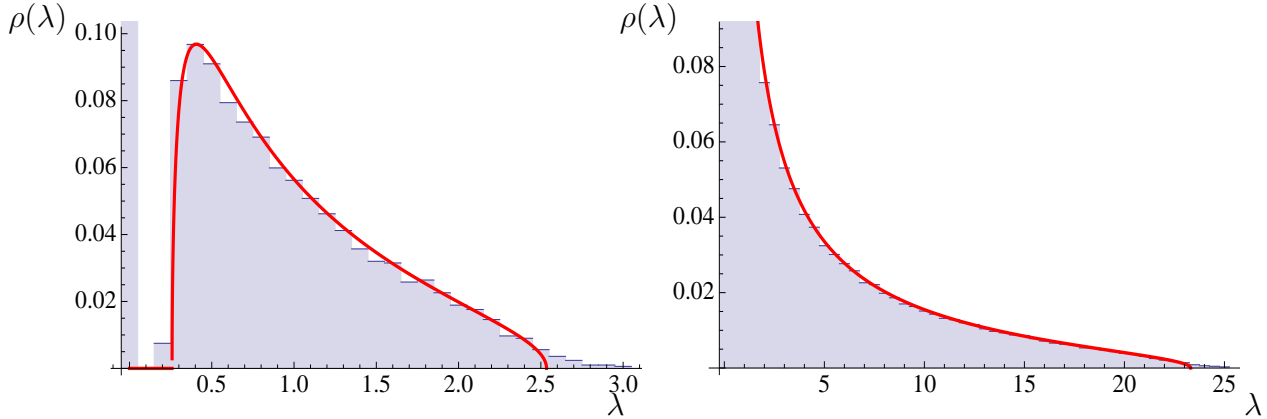
and

$$\lambda_{\pm} = \frac{1}{27} \left( \alpha(\alpha(2\alpha - 9) + 108) \pm 2\sqrt{(\alpha - 9)^2\alpha(\alpha + 3)^3 + 27} \right).$$

Here somewhat the situation is reverse, namely, there is a gap for  $\alpha < 1$  and no gap for  $\alpha \geq 1$ . In figure 7 we show some plots of our theoretical predictions and comparisons with numerical diagonalisation.

Finally, for small  $\alpha$  we find again that the continuous part is proportional to the Wigner law after doing the change of variables  $\lambda = 2\sqrt{3\alpha}w + 1$ .

We finish by pointing out that the cases  $\alpha^{(1)}$  and  $\alpha^{(3)}$  are related by  $(\lambda, \alpha) \rightarrow (\lambda/\alpha, 1/\alpha)$ .



**Figure 7.** Case  $\alpha^{(3)} = (1, 1, \alpha)$  for  $\alpha = 1/10$  (for the histograms we have used  $N(1) = N(2) = N(3) = 100$  and  $N(4) = 10$  averaged over 1000 samples) with Dirac delta and  $\alpha = 3$  (for the histograms we have used  $N(1) = N(2) = N(3) = 100$  and  $N(4) = 300$  averaged over 1000 samples).

## 5. Three observations and some exact results for general $M$

### 5.1. Three numerical observations

Based on the detailed derivations we have done for  $M = 1, 2$  and  $3$  and on some, admittedly rather weak, numerical experiments for larger values of  $M$ , we would like to point out three observations that seem to be correct. Our first observation is:

**Observation 1 (The weight of the Dirac delta)** *For any  $M$  the weight associated with the Dirac delta is*

$$\omega(\alpha) = 1 - \min(1, \alpha_1, \dots, \alpha_M).$$

This is obviously true for  $M = 1, 2$ , and  $3$  and seems to be true numerically for  $M \geq 4$ . However it is quite challenging to rigorously check this numerically for two reasons: firstly, for any value of  $M$  (let us say moderately big) the error associated with the numerical diagonalisation will not provide exactly zero eigenvalues but instead very small ones. If there were to be a gap between the zero -or very small- eigenvalues to the continuous spectrum, an estimate of the weight is certainly possible. However, in many cases there is either no gap or the gap is very small (in the sense that it is numerically indistinguishable as not having gap). We then run into the situation that the continuous part is mixed with the Dirac delta, not a pleasant situation to have numerically. A way around this is to calculate the cumulative probability function and then look at its derivative at zero. However, to have a good estimate for its derivative, the number of samples and the size of matrices must be large enough.

The second observation has to do with the lack of gap (but not necessarily of Dirac delta) when the ensemble of matrices is such that the first and the last matrices in  $X(1) \cdots X(M)$  are square for the family of parameters  $\alpha^{(s)}$ :



**Observation 2 (No gap)** *The continuous part of spectral density for the family of parameters  $\alpha^{(s)}$  for  $s = 2, \dots, M-1$  has no gap in the spectral density for any value of the parameter  $\alpha$ .*

Again we would like to check somewhat carefully this observation numerically, but we find ourselves again in a difficult situation trying to disentangle the contribution of the Dirac delta from the continuous part.

The final observation has to do with the defined shape the spectral density takes for small  $\alpha$  for the family of spectral densities with parameters  $\alpha^{(s)}$ :

**Observation 3 (Asymptotic shape)** *Consider the spectral density  $\rho_M(\lambda|\alpha^{(s)})$  for  $s = 1, \dots, M$ . Then for small  $\alpha$  and for  $s = 1, \dots, M-1$  one finds that*

$$\rho_M(\lambda|\alpha^{(s)}) = (1 - \alpha)\delta(w) + \alpha\rho_{M-s}(\lambda|\{\alpha_i = 1\}_{i=1}^{M-s}) + \mathcal{O}(\alpha^2)$$

*while for  $s = M$  one finds that the asymptotic behaviour of the continuous part is given by the Wigner semi-circle law  $\frac{2}{\pi}\sqrt{1-w^2}$  after rescaling to the new variable  $\lambda = 2\sqrt{M\alpha}w + 1$ .*

As we have seen this is certainly true for the exact cases we have discussed before. For  $M \geq 4$  we would like to check this at least numerically. However we run into another impasse: to avoid finite size effects we must work with large matrix sizes with small ratio  $\alpha$ . We could of course achieve small values of  $\alpha$  using small matrices but then the finite size effects -not captured in this approach- are too significant to check carefully this observation.

A mathematically grounded explanation of these observations can be found in [3, 18].

## 5.2. Endpoints of the continuous part of the spectrum

Starting from the polynomial equation (16), it is possible to derive equations for the endpoints of the support of the spectral density (this simple trick can also be found in [3]). Indeed, doing the derivative of  $v$  with respect to  $z$  and noticing that  $dv/dz$  is not well-defined at the endpoints, we obtain the following condition for  $v(z_B)$  at the boundary:

$$\frac{\alpha_M v(z_B)}{1 + \alpha_M v(z_B)} + \sum_{\ell=1}^M \frac{\alpha_\ell v(z_B)}{\alpha_\ell + \alpha_M v(z_B)} = 1. \quad (19)$$

We obviously cannot solve exactly this equation for any set of parameters  $\alpha$ , but we can for the family of parameters used before.

**5.2.1. Square matrices** Here  $\alpha = (1, \dots, 1)$  and the eq. (19) gives  $v(z_B) = 1/M$ , which after inserting it in the polynomial equation (16) yields the upper endpoint

$$\lambda_+ = M \left( 1 + \frac{1}{M} \right)^{M+1}. \quad (20)$$

Eq. (20) agrees with our previous ones for  $M \leq 3$ . We know that the lower endpoint is zero -as there is no gap- but we are not able to get  $\lambda_-$  by starting from all square matrices, as we will see below.

*5.2.2. Family of parameters  $\alpha^{(s)}$*  For this family of parameters, the equation (19) becomes

$$s \frac{\alpha v(z_B)}{1 + \alpha v(z_B)} + (M - s + 1) \frac{v(z_B)}{1 + v(z_B)} = 1.$$

Solving it and plugging the solution back into the polynomial equation, provides the following result:

$$z_{\tau,B}(\alpha, M, s) = \frac{1}{\Delta_\tau(\alpha, M, s)} (1 + \alpha \Delta_\tau(\alpha, M, s))^s (1 + \Delta_\tau(\alpha, M, s))^{M-s+1},$$

for  $\tau \in \{-1, 1\}$  and for  $s = 1, \dots, M$ , and where we have defined

$$\Delta_\tau(\alpha, M, s) = \frac{-\alpha(s-1) - M + s + \tau \sqrt{((\alpha-1)(s-1) + M-1)^2 + 4\alpha M}}{2\alpha M}.$$

A similar formula can be found in [3]. One should first note that for  $\alpha = 1$  we recover back eq. (20) plus the lower endpoint at zero. In the general case, however,  $z_{\tau,B}(\alpha, M, s)$  gives the upper endpoint ( $\tau = 1$ ) correctly, while for the lower endpoint we must tweak the formula \*. The correct lower and upper endpoints turn out to be:

$$\lambda_-(\alpha, M, s) = z_{-,B}(\alpha, M, s) [\delta_{s,1} \Theta(\alpha-1) + \delta_{s,M} \Theta(1-\alpha)], \quad \lambda_+(\alpha, M, s) = z_{+,B}(\alpha, M, s).$$

*5.2.3. Exact expression of the spectral density for square matrices* Finally, it is possible to find an exact expression for the spectral density for any  $M$  and square matrices by using the method of series inversion. Indeed, after changing variables  $u = 1 + v$ , the polynomial equation (16) can be written as follows:

$$t = \frac{u}{\phi(u)}, \quad \phi(u) \equiv (u-1)^{\frac{1}{M+1}}, \quad t = z^{\frac{1}{M+1}} e^{\frac{2\pi k i}{M+1}}, \quad k = 0, \dots, M.$$

The idea is then to obtain  $u(t)$  by inverting the series. We do this using Lagrange-Burmänn formula:  $u(t) = \sum_{n=1}^{\infty} \frac{1}{n} [w^{n-1}] \phi^n(w) t^n$ , where here the symbol  $[z^n]f(z)$  is an operator that gives the  $n$ -th coefficient of the power series of  $f(z)$ . After some algebra, going back to the variable  $\sigma(1)$ , choosing the  $k$  that gives the spectral density (it turns out to be  $k = 0$ ) we end up with

$$\rho_M(\lambda) = \frac{1}{\pi} \left\{ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \binom{\frac{n}{M+1}}{n-1} \lambda^{\frac{n}{M+1}-1} \sin \left( \frac{\pi n}{M+1} \right) \right\} \mathbf{1}_{\lambda \in [0, M(1+\frac{1}{M})^{M+1}]}.$$

This result, which is also called Fuss-Catalan distribution, was also derived in [19]. Let us check that indeed we recover the previous cases:

\* This is what we also found in our analysis of  $M = 3$ , in which the endpoints of spectral density are related to the roots of  $P_4$ . Actually, one can see that for  $M = 3$ ,  $z_{\tau,B}(\alpha, M, s)$  provides two of the roots of  $P_4$ .

- For  $M = 1$  we have

$$\rho_1(w) = \frac{1}{2\pi} \sqrt{\frac{4-w}{w}}. \quad (21)$$

- For  $M = 2$  we have

$$\rho_2(w) = \frac{1}{2\sqrt{3}\pi w} \left( w^{2/3} {}_2F_1 \left[ \frac{1}{6}, \frac{2}{3}; \frac{4}{3}; \frac{4w}{27} \right] - 3w^{1/3} {}_2F_1 \left( -\frac{1}{6}, \frac{2}{3}; \frac{2}{3}; \frac{4w}{27} \right) \right]$$

which, after expressing the hypergeometric functions using basic ones, we obtain the results obtained previously.

- For  $M = 3$  we obtain a rather involved expression in terms of generalised hypergeometric functions:

$$\begin{aligned} \rho_3(w) = & \frac{1}{65536\pi w\sqrt{2}} \left( 65536\sqrt[4]{w} {}_6F_5 \left( -\frac{1}{24}, \frac{1}{8}, \frac{7}{24}, \frac{11}{24}, \frac{5}{8}, \frac{19}{24}; \frac{1}{4}, \frac{3}{8}, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}; \frac{729w^2}{65536} \right) \right. \\ & - 16384\sqrt{2}\sqrt{w} {}_6F_5 \left( \frac{1}{12}, \frac{1}{4}, \frac{5}{12}, \frac{7}{12}, \frac{3}{4}, \frac{11}{12}; \frac{3}{8}, \frac{1}{2}, \frac{5}{8}, \frac{7}{8}, \frac{9}{8}; \frac{729w^2}{65536} \right) \\ & - 128\sqrt{2}w^{3/2} {}_6F_5 \left( \frac{7}{12}, \frac{3}{4}, \frac{11}{12}, \frac{13}{12}, \frac{5}{4}, \frac{17}{12}; \frac{7}{8}, \frac{9}{8}, \frac{11}{8}, \frac{3}{2}, \frac{13}{8}; \frac{729w^2}{65536} \right) \\ & - 2048w^{3/4} {}_6F_5 \left( \frac{5}{24}, \frac{3}{8}, \frac{13}{24}, \frac{17}{24}, \frac{7}{8}, \frac{25}{24}; \frac{1}{2}, \frac{5}{8}, \frac{3}{4}, \frac{9}{8}, \frac{5}{4}; \frac{729w^2}{65536} \right) \\ & - 224w^{5/4} {}_6F_5 \left( \frac{11}{24}, \frac{5}{8}, \frac{19}{24}, \frac{23}{24}, \frac{9}{8}, \frac{31}{24}; \frac{3}{4}, \frac{7}{8}, \frac{5}{4}, \frac{11}{8}, \frac{3}{2}; \frac{729w^2}{65536} \right) \\ & \left. - 39w^{7/4} {}_6F_5 \left( \frac{17}{24}, \frac{7}{8}, \frac{25}{24}, \frac{29}{24}, \frac{11}{8}, \frac{37}{24}; \frac{9}{8}, \frac{5}{4}, \frac{3}{2}, \frac{13}{8}, \frac{7}{4}; \frac{729w^2}{65536} \right) \right). \end{aligned}$$

It can be checked that this again agrees with the expression found in the previous section.

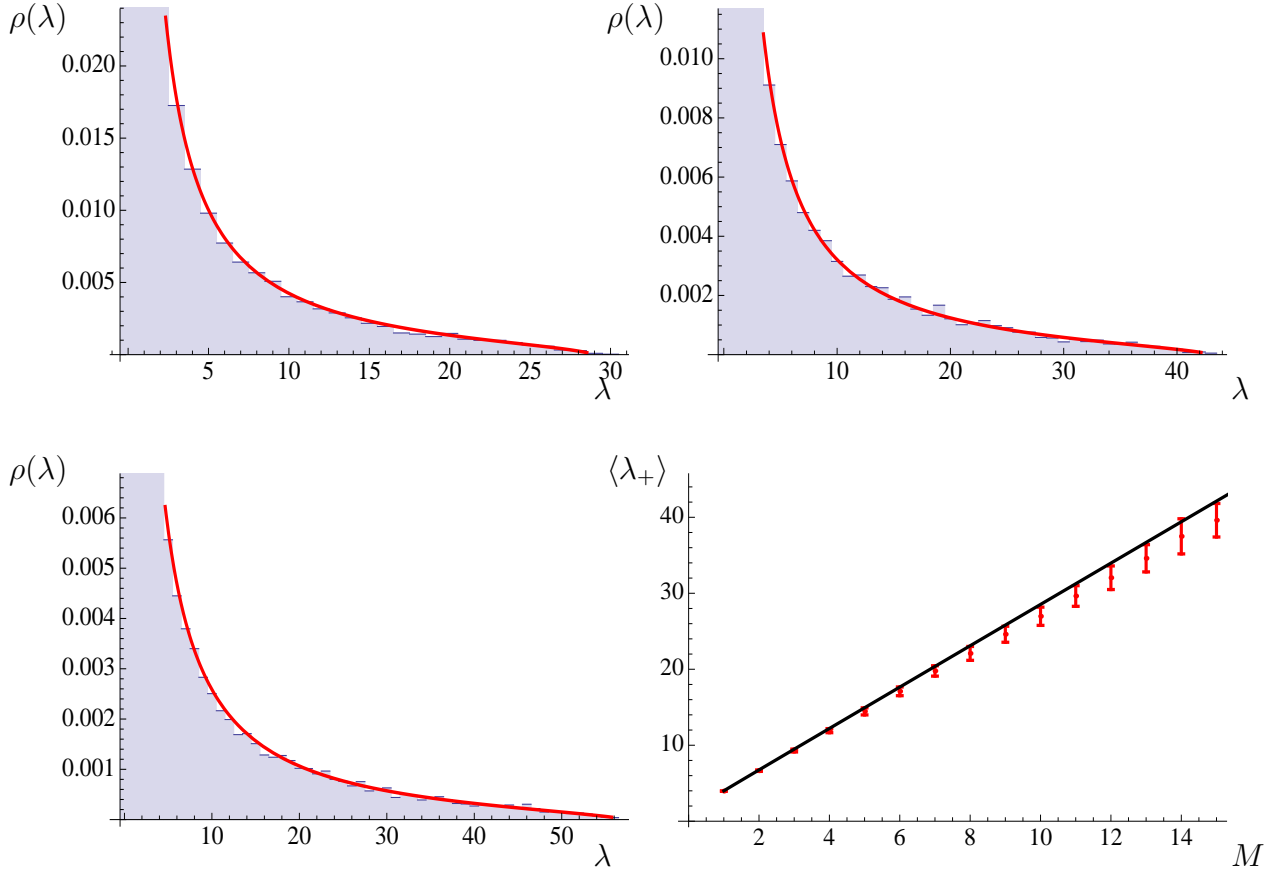
It is also worth noticing that given a generalized Hypergeometric function  ${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; z)$  with  $p = q + 1$ , the series converges for  $|z| \leq 1$ . This completely agrees with the formula (20) of the upper endpoint we found previously: for  $M = 2$  we have that  $|4w/27|$  or  $|w| \leq 27/4$ , while for  $M = 3$   $|729w^2/65536|$  yielding  $|w| \leq 256/27$ .

In figure 8, we have plotted our theoretical findings for  $M = 10, 15$  and  $20$  for square matrices and compared them with results from numerical diagonalisation. Moreover, we have numerically checked the formula (20) for the upper endpoint.

## 6. Summary, part II and other future work

In this work we have tackled the problem of the spectral density of products of Wishart diluted random matrices using the cavity method. We have focused in this part in analysing the dense limit. This has allowed us to make contact with already known results and to bring to light new ones.

In part II we will show that a polynomial equation similar to eq. (16) can also be derived for random regular graphs. This allows us to obtain exact results for  $M$  up to



**Figure 8.** Spectral density for square matrices for  $M = 10, 15$ , and  $20$  (left to right, up to bottom). Red line plots correspond to the theoretical prediction of formula (21)). For the numerical diagonalisation, we have used square matrices of size 1000 and averaged over 1000 samples. In the last plot we have estimated numerically the upper endpoint using square matrices of size 1000 and averaging over a 100 samples. We have also plotted the error bars indicating the typical fluctuations of the largest eigenvalue. As we can see, as  $M$  increases the typical fluctuation also increases. This is due to the fact that the fluctuations are amplified when the number of matrix multiplications increases.

three and also permits some humble exact results for general values of  $M$ . For general diluted graphs, the spectral density can only be obtained by solving the self-consistency equations numerically by using either population dynamics or belief-propagation.

Apart of the content of part II, there are other research lines that we are currently persuing. As an instance, we are exploring in which way correlation between matrices impacts the statistical properties derived from this type of ensembles [11], and extensions of the present work to products of non-Hermitian diluted matrices [20].

## Acknowledgments

IPC would like to thank Jesper Ipsen for correcting our second numerical observation and pointing out to relevant literature, making this a better rounded paper.

- [1] Burda Z, Janik R A and Waclaw B 2010 *Phys. Rev. E* **81**(4) 041132
- [2] Burda Z, Jarosz A, Livan G, Nowak M A and Swiech A 2010 *Phys. Rev. E* **82**(6) 061114
- [3] Akemann G, Ipsen J R and Kieburg M 2013 *Phys. Rev. E* **88**(5) 052118
- [4] Akemann G, Kieburg M and Wei L 2013 *J. Phys. A* **46** 275205
- [5] Akemann G, Burda Z, Kieburg M and Nagao T 2013 *preprint arXiv:1310.6395*
- [6] Gudowska-Nowak E, Janik R A, Jurkiewicz J and Nowak M A 2003 *Nucl. Phys. B* **670** 479 – 507
- [7] Götze F and Tikhomirov A 2010 *preprint arXiv:1012.2710*
- [8] Andrea Crisanti Giovanni Paladin A V 1993 *Products of Random Matrices in Statistical Physics* vol 104 (Springer Series in Solid-State Sciences)
- [9] Edwards S and Jones R C 1976 *Journal of Physics A: Mathematical and General* **9** 1595
- [10] Vinayak 2013 *Phys. Rev. E* **88**(4) 042130
- [11] Pérez Castillo I 2014 *in preparation*
- [12] Mezard M and Montanari A 2009 *Information, physics, and computation* (Oxford University Press)
- [13] Rogers T, Pérez Castillo I, Kühn R and Takeda K 2008 *Phys. Rev. E* **78**(3) 031116
- [14] Rogers T and Pérez Castillo I 2009 *Phys. Rev. E* **79**(1) 012101
- [15] Shamir M and Sompolinsky H 2000 *Physical Review E* **61** 1839
- [16] Metz F L, Parisi G and Leuzzi L 2014 *arXiv preprint arXiv:1403.2582*
- [17] Rogers T, Pérez Vicente C, Takeda K and Pérez Castillo I 2010 *J. Phys. A* **43** 195002
- [18] Ipsen J R and Kieburg M 2013 *arXiv preprint arXiv:1310.4154*
- [19] Penson K A and Życzkowski K 2011 *Physical Review E* **83** 061118
- [20] Pérez Castillo I 2014 *in preparation*

## Appendix A. Derivation of polynomial equation

In the cavity equations we introduce the new variables  $u(t) = \sigma(t)/(J^2(t-1)\sigma(t-1))$  and  $v(t) = \sigma(t)\Sigma(t)$ . Upon also defining  $\sigma(0) = 1/(zJ^2(0))$  we obtain the following equations:

$$\begin{aligned} v(t) &= \frac{\alpha_t}{\alpha_{t-1}} (1 - u(t+1)) \text{ with } t = 1, \dots, M-1 \\ v(t) + u(t) &= 1 \text{ with } t = 1, \dots, M \\ \Sigma(M) &= -\frac{\alpha_M}{\alpha_{M-1}} \frac{J^2(M)}{1 - J^2(M)\sigma(M)} \end{aligned}$$

Using these new variables the derivation of the polynomial equation (16) becomes fairly simple:  $u(t)$  gives a recursion for  $\sigma(t)$ , whose solution is at  $t = M$ :

$$\sigma(M) = \sigma(0) \prod_{s=1}^M J^2(s-1)u(s) \quad (\text{A.1})$$

The solutions for  $v(t)$  and  $u(t)$  obviously are

$$v(t) = \frac{v(1)}{\alpha_{t-1}}, \quad u(t) = 1 - \frac{1 - u(1)}{\alpha_{t-1}}$$

Moreover for  $\sigma(M)$  we find that

$$\sigma(M) = \frac{1 - u(1)}{J^2(M)\alpha_M \left[ \frac{1-u(1)}{\alpha_M} - 1 \right]}$$

Using all these results back into eq. (A.1) we write

$$\frac{1 - u(1)}{J^2(M)\alpha_M \left[ \frac{1-u(1)}{\alpha_M} - 1 \right]} = \sigma(0) \prod_{s=1}^M J^2(s-1)u(s)$$

After arranging terms and recalling that  $u(1) = z\sigma(1)$  and  $\sigma(0) = 1/(zJ^2(0))$  we finally obtain

$$\mathcal{J}_M^2 \prod_{s=1}^{M+1} \left( 1 + \frac{\alpha_M}{\alpha_{s-1}} v \right) = vz, \quad v \equiv -\frac{1 - z\sigma(1)}{\alpha_M}, \quad \mathcal{J}_M \equiv \prod_{s=1}^M J(s)$$

## Appendix B. Case $M = 2$

In this case we have the following cubic polynomial equation  $a\sigma^3(1) + b\sigma^2(1) + c\sigma(1) + d = 0$  with coefficients:

$$d = \frac{z}{\alpha_2}, \quad c = \frac{z(-\alpha_1 z + (1 - \alpha_1)(1 - \alpha_2))}{\alpha_1 \alpha_2}, \quad b = \frac{z^2(\alpha_1 + \alpha_2 - 2)}{\alpha_1 \alpha_2}, \quad a = \frac{z^3}{\alpha_1 \alpha_2}. \quad (\text{B.1})$$

There are various ways to write down the solution to a cubic equation. Here we choose the standard one and write the three solutions as  $x_k = -\frac{1}{3a} \left( b + u_k C + \frac{\Delta_0}{u_k C} \right)$  for  $k = 1, 2$  and  $3$ , and where  $u_1 = 1$ ,  $u_2 = \frac{-1+i\sqrt{3}}{2}$  and  $u_3 = \frac{-1-i\sqrt{3}}{2}$  and where we have defined:

$$C = \sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}}, \quad \Delta_0 = b^2 - 3ac, \quad \Delta_1 = 2b^3 - 9abc + 27a^2d. \quad (\text{B.2})$$

After some tedious algebra one can cancel some factors yielding

$$x_k = -\frac{1}{3z} \left( \gamma - 3 + u_k \tilde{C} + \frac{\tilde{\Delta}_0}{u_k \tilde{C}} \right)$$

with

$$\begin{aligned} \tilde{C} &= \sqrt[3]{\frac{\tilde{\Delta}_1 + \sqrt{\tilde{\Delta}_1^2 - 4\tilde{\Delta}_0^3}}{2}}, \quad \tilde{\Delta}_0 = 3z\alpha_1 + \gamma^2 + 3(1 - \gamma_1\gamma_2) \\ \tilde{\Delta}_1 &= (2\gamma^3 - 9(\alpha_1\gamma_1 + \alpha_2\gamma_2 + \alpha_1\alpha_2(\gamma - 1)) + 9z\alpha_1\gamma), \end{aligned}$$

where we have defined  $\gamma = 1 + \alpha_1 + \alpha_2$  and  $\gamma_i = 1 + \alpha_i$ . It is important to notice that factor  $1/z$  in the expression of  $x_k$  allows us to use distribution theory comfortably to get the continuous part of the spectral density as well as the weight of Dirac delta. Now, a numerical check reveals that the spectral density is given by  $k = 3$ , while the endpoints are given by the roots of  $\tilde{\Delta}_1^2 - 4\tilde{\Delta}_0^3 = 0$ . It turns out that the latter is a cubic polynomial equation  $az^3 + bz^2 + cz + d = 0$  with coefficients

$$\begin{aligned} d &= -27(-1 + \alpha_1)^2(\alpha_1 - \alpha_2)^2(-1 + \alpha_2)^2, \\ c &= -216\alpha_1(1 + \alpha_1 + \alpha_2)^2(1 - (1 + \alpha_1)(1 + \alpha_2)) - 324\alpha_1(1 - (1 + \alpha_1)(1 + \alpha_2))^2 \\ &\quad - 162\alpha_1(1 + \alpha_1 + \alpha_2)(\alpha_1(1 + \alpha_1) + \alpha_2(1 + \alpha_2) + \alpha_1\alpha_2(\alpha_1 + \alpha_2)), \\ b &= -27\alpha_1^2(1 + \alpha_1 + \alpha_2)^2 - 324\alpha_1^2(1 - (1 + \alpha_1)(1 + \alpha_2)), \\ a &= -108\alpha_1^3. \end{aligned}$$

### Appendix C. Case $M = 3$

In this case we write the quartic polynomial equation as  $a\sigma^4(1) + b\sigma^3(1) + c\sigma^2(1) + d\sigma(1) + e = 0$  with coefficients

$$\begin{aligned} a &= \frac{\lambda^4}{\alpha_1\alpha_2\alpha_3}, \quad b = \frac{(-3 + \alpha_1 + \alpha_2 + \alpha_3)\lambda^3}{\alpha_1\alpha_2\alpha_3}, \\ c &= \frac{(3 + \alpha_2(-2 + \alpha_3) - 2\alpha_3 + \alpha_1(-2 + \alpha_2 + \alpha_3))\lambda^2}{\alpha_1\alpha_2\alpha_3}, \\ d &= \frac{\lambda((-1 + \alpha_1)(-1 + \alpha_2)(-1 + \alpha_3) - \alpha_1\alpha_2w)}{\alpha_1\alpha_2\alpha_3}, \quad e = \frac{\lambda}{\alpha_3}. \end{aligned}$$

Using the standard expression for the four solutions one ends up with the following expression:

$$\begin{aligned} x_{\sigma_1, \sigma_2} &= -\frac{-4 + \gamma}{4\lambda} - \sigma_1 \frac{1}{\lambda} \mathcal{S}(\lambda) + \sigma_2 \frac{1}{2\lambda} \sqrt{-4\mathcal{S}^2(\lambda) - \frac{\gamma^2 - 4\gamma_2}{4} + \sigma_1 \frac{1}{8} \frac{Q_1(\lambda)}{\mathcal{S}(\lambda)}} \\ \mathcal{S}(\lambda) &= \frac{1}{2} \sqrt{-\frac{\gamma^2 - 4\gamma_2}{12} + \frac{1}{3} \left( \sqrt[3]{\frac{P_2(\lambda) + \sqrt{P_4(\lambda)}}{2}} + \frac{P_1(\lambda)}{\sqrt[3]{\frac{P_2(\lambda) + \sqrt{P_4(\lambda)}}{2}}} \right)} \end{aligned}$$

with  $\sigma_1, \sigma_2 \in \{-1, 1\}$  and  $\gamma = 1 + \alpha_1 + \alpha_2 + \alpha_3$   $\gamma_2 = 1 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2$ . Here  $Q_1(\lambda)$ ,  $P_1(\lambda)$ ,  $P_2(\lambda)$ , and  $P_4(\lambda)$  are the following polynomials

$$\begin{aligned}
Q_1(\lambda) &= \alpha_1^3 + (-1 + \alpha_2 - \alpha_3)(1 + \alpha_2 - \alpha_3)(-1 + \alpha_2 + \alpha_3) - \alpha_1^2(1 + \alpha_2 + \alpha_3) \\
&\quad - \alpha_1(\alpha_2^2 + (-1 + \alpha_3)^2 - 2\alpha_2(1 + \alpha_3 - 4\lambda)) , \\
P_1(\lambda) &= \alpha_3^2 - \alpha_2\alpha_3(1 + \alpha_3) + \alpha_2^2(1 + (-1 + \alpha_3)\alpha_3) - \alpha_1(\alpha_2(1 + \alpha_2) + \alpha_3 \\
&\quad + (-6 + \alpha_2)\alpha_2\alpha_3 + (1 + \alpha_2)\alpha_3^2) \\
&\quad + \alpha_1^2(1 + \alpha_2^2 + (-1 + \alpha_3)\alpha_3 - \alpha_2(1 + \alpha_3 - 3\lambda)) + 3\alpha_1\alpha_2(1 + \alpha_2 + \alpha_3)\lambda , \\
P_2(\lambda) &= (\alpha_2(-2 + \alpha_3) + \alpha_3)(\alpha_2 + (-2 + \alpha_2)\alpha_3)(-\alpha_3 + \alpha_2(-1 + 2\alpha_3)) \\
&\quad - 3\alpha_1^2(\alpha_3 + (-4 + \alpha_3)\alpha_3^2 + \alpha_3^3(1 + \alpha_3)) \\
&\quad + 2\alpha_2^2(-2 + \alpha_3 - 2\alpha_3^2) + \alpha_2(1 + \alpha_3)(1 + \alpha_3 + \alpha_3^2)) \\
&\quad + 9\alpha_1^2\alpha_2(1 + \alpha_2^2 + (-3 + \alpha_3)\alpha_3 - 3\alpha_2(1 + \alpha_3))\lambda \\
&\quad + 27\alpha_1^2\alpha_2^2\lambda^2 + \alpha_1^3(2\alpha_2^3 + (-2 + \alpha_3)(1 + \alpha_3)(-1 + 2\alpha_3) - 3\alpha_2(1 + (-4 + \alpha_3)\alpha_3) \\
&\quad - 3\alpha_2^2(1 + \alpha_3 - 3\lambda) + 9\alpha_2(1 + \alpha_3)\lambda) \\
&\quad + 3\alpha_1(-\alpha_3^2(1 + \alpha_3) + \alpha_2^3(-1 + 3\lambda + \alpha_3(4 - \alpha_3 + 3\lambda)) \\
&\quad + \alpha_2\alpha_3(4 + 3\lambda + \alpha_3(-2 + 4\alpha_3 + 3\lambda)) \\
&\quad - \alpha_2^2(1 - 3\lambda + \alpha_3(2 + \alpha_3(2 + \alpha_3 - 3\lambda) + 9\lambda))) , \\
P_4(\lambda) &= P_2^2(\lambda) - 4P_1^3(\lambda) .
\end{aligned}$$

A numerical check reveals that the spectral density is given by the choice  $\{\sigma_1, \sigma_2\} = \{-1, 1\}$  leading to the reported expression of the spectral density for  $M = 3$ .